


```

FEATURES
    source          1..18
                   /organism="unknown"
BASE COUNT      2 a      5 c      7 g      4 t
ORIGIN
Query Match      100.0%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      1 GTGGTGGTCCAGCCGTA 18
        |||
        1 GTGGTGGTCCAGCCGTA 18

RESULT 2
AR098113/C
LOCUS      AR098113      218 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION Sequence 21 from patent US 6074832.
ACCESSION  AR098113
VERSION     AR098113.1 GI:12807370
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 218):
            Venter, P.J., Brewer, G.J., Yuzbasyan-Gurkan, V., Schall, W.D. and
            Dufrenoy, J.
            DNA encoding canine von Willebrand factor and methods of use
            Patent: US 6074832 A 21 13-11-2000;
            Location/Qualifiers
            source          1..218
                           /organism="unknown"
BASE COUNT      28 a      67 c      74 g      49 t
ORIGIN
Query Match      100.0%; Score 18; DB 6; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      1 GTGGTGGTCCAGCCGTA 18
        |||
        1 GTGGTGGTCCAGCCGTA 190

RESULT 3
H00WVFA04/C
LOCUS      H00WVFA04      734 bp      DNA      linear      PRI 14-JAN-1995
DEFINITION Human von Willebrand factor gene, exon 7.
ACCESSION  M25831 M25716
VERSION     M25831.1 GI:340320
KEYWORDS    coagulation factor VIII; von Willebrand factor
SEGMENT     4 of 38
            Human placenta, leukocyte, fetal liver and cell line EB19 DNA
            clones lambda: [1,2,7,9,11], h[6,7,18], c[1,2,3,4,5,9,14,18] and
            pW[5,8,12,3].
            Homo sapiens.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 734):
            Mancuso, D., Tiller, P. A., Westfield, A., Worrall, N. K.,
            Shelton-Inloes, B. B., Sorace, J. M., Alevis, Y. G. and Sadler, J. E.
            Structure of the gene for human von Willebrand factor
            J. Biol. Chem. 264 (33), 19514-19527 (1989)
JOURNAL     90062044
MEDLINE     258182
PUBMED
COMMENT     Draft entry and computer-readable sequence for [1] kindly provided
            by J.E.Sadler, 30-JUN-1989.
FEATURES
    source          1..734
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /map="12p11p12"

```

```

        intron      -1..108
                   /gene="VWF"
                   /number=6
        exon         109..125
                   /gene="VWF"
                   /number=7
        intron       126..1734
                   /gene="VWF"
                   /number=7
BASE COUNT      106 a      247 c      216 g      165 t
ORIGIN
Query Match      100.0%; Score 18; DB 9; Length 734;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY      1 GTGGTGGTCCAGCCGTA 18
        |||
        1 GTGGTGGTCCAGCCGTA 298

RESULT 4
H00WVFA04/C
LOCUS      H00WVFA04      4429 bp      RNA      linear      PRI 21-MAR-1995
DEFINITION Human mRNA fragment (5' terminus) for von Willebrand factor (VWF).
ACCESSION  X04146
VERSION     X04146.1 GI:37941
KEYWORDS    Glycoprotein; plasma protein; signal peptide; von Willebrand
            factor.
SOURCE      Homo sapiens.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 4429):
            Verweij, C. L., Diegenard, P. J., Hart, M. and Pannekoek, H.
            Full-length von Willebrand factor (VWF) cDNA encodes a highly
            repetitive protein, considerably larger than the mature VWF subunit
            EMBO J. 5 (8), 1839-1847 (1986)
JOURNAL     87004509
MEDLINE     3019665
PUBMED
COMMENT     <H00WVFA04> corresponding to pos. 2458-2849 and
            (M10320) corresponding to pos. 4894-4429.
            Location/Qualifiers
            source
            CDS
            1..4429
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            439..4429
            /note="Protein VWF (aa 22 to 1378)"
            /codon_start=1
            /protein_id="CAA27765.1"
            /db_xref="GI:37942"
            /translation="MTTPAPPAVITLAIATIIPTGTCAPPTGRPSSTARGSLFSGHFN
            FPDGSMVSFAGVCSYLLAGGQKRSSTIDPQNCQKRSISVYLDFPDIHLFVNTV
            FLDGKLVSMVASKGLYILEAVYVLSLHAYFVVAHILGSLNFVLSLQYFNRG
            LCGNCFEAEIDEMTORGTTISDPYFANSWALSSGQWCPKASPESSGNSGEMO
            KGLMEQVLLKSTSVLAKCHLVDPPEFVALSKELCHLHAGELACALLFYATTA
            QEGVNLGMDTHSACSVPCAGMEYQCSFKARTQSHINEMQDERVQKSTPRN
            QULDEGLVRSFIPVHSKRPVPSISLSPENITLPSNWTISNEEDPDECLATG
            QSHRSFSDNERFTFSGLQVYLARDQDHSFSLVETVCAADKRAVACRSVTVKGLG
            ILNSIVLKLRKAGVAMQGVVQLDLKGLPELQVITAVAPVSYCEELQMWVQPRLE
            LVKLSVYACKTGICNGNNGGDEFTFSGIAGREYVAPGEIVKIGNTGCVHPRKW
            SDPALNPRMRESSEACAVLTSPTEFCRAVAVSPPLRYKRRVYVSGSLQKRETLG
            ALASYAACKRGVAVVAPFPPQCLNCRKGLVYVQVHINLTKSLSYPERNEA
            CLEGGPCPGIYMDERGVCPKACQPDYDGLPEDEI FSDHHTMCTEDGMRITM
            SGVSGSLPQAVLSSTSHRSKRSLSRPPMKVTPAINLAFLELTKIQUNLILE
            QMSMGVSSGLCPGQVGRHRCVLRCPHCKEYVAPGEIVKIGNTGCVHPRKW
            NCPHVCDAICSTIGMAHYLTPDGLKYLTPGKQVYLVDQGGSNQPRILLVKKGG
            SHPSVCKKRVTLIVEGELELPDGEYVNNKRPKQETHEEVESPPYITLLGKALSV
            VMDPHLSTSVLKQTVGRKVGLQGNFSTGNNQNTISNIVGFPPVPRNGVAVSSQ
            CADTRKVPDLSPPATNNIMKQIVSSSRITSDVPYCPKRLVPEYDLVLTVDI
            QDSKSLDCAQCHLQILAAVAVHVAQMDKVLIMFALVLPQSLPEFNLKENVLCKWKY
            NSZAVACQVITQGHETLAVVGVDSGHMCPKRLDRLDGLQVAFEDQVYGVKQK

```



```

/codon_start=1
/protein_id="CAA27972.1"
/db_xref="GI:37947"
/db_xref="SMISS-PROT:P04275"
/translation="MTPARFAGVILALILILPSTLAEGLFCHSLAAQLNLSNIFVN
TFDGSVAFAGVSYLLAGGQKRPSTIGDPQGRKPSITLSVYLGPFDHLPFVNGT
TQGDQRMVPSYAKQVLETFAGYKILSGAYAGVVARIDGSGVQVLLSDRYNKG
LCGNENI:PAEDPMTUORGLTSDYADNSWALSGRPMERAPSPSSNTISGKMG
KGLMEUOLLKSTSVARCPHLVDPEPVALCEITLCEGLNEMLEKRYGVSJYEDS
OGLDGLVSTSTECPCVHSKRPRPGTSLSTLNTLITNSLMITSNFEPGFLVTS
OSHSKSPDNRTFTFSGICVLLARDQDHSFSTIVTQVADRDADACSTRSVIRL
LHNSLVKLKHGAVAMQDQVLLKGLDRLTQTVLAVSLYSYGBLDMWIKRRL
LVKLSVYAKRTGICGNVNGNIGFILTSGJAEPEDEFGAMKILHQQQLQJH
SDPALNRPMTFSEEAVALTSPFACBPAPVLELKNCRYPDVASJYGEJGQ
ALASYAAGAGRVAVAREGCELCNPGYVLOGCPGNCITGSLSTYDEGCEMA
CLEGFCPCGLYMDRGDVCVACQCPYVGGELFQPDITFSDHHTXGCEDEGMHT
SGVCSLLPDAVLSSPLSHRSKSLSCRPVAVKVCADMLRAKGLBECTKQCNTOLE
CTMGQVSGGLCPPMVPHENRVALFPFPGHQRKYEAGETVKIGCNTGVDRKW
NCTDHVTAQSTIGMAHYITFDQIKYLPFPGVYLVQYVQSGNSMPTLILVGNKG
SHPSYKCKKRVITIVGGEIELPDGEVAVRPMKDEHFEEVSGEYIILLKSLAL
VMDRLISVYLKQYGEKVLGNGDLGNLJLSNVLEDEDPYDEGNSKRVSSQ
CADTRKVPDSSPATCNHNIKOTMVSOSRILTSDEPOCNKLMDPEPVDGCTIT
CSCESTGFCFCTIAYAHVACGKVKYIKTALCPKSTERRNIRHNYTTPMY
NSCAPQVYQGHHPILACPYQVYSGHARQPEKILDELIVQVDEDPVAVGR
RPSAKKATLNPDPHQICHDVANIJCACQDEGLVAVPDAVSPPTLVYEDI
SEPLHDFVSRLLDLVELDSSSLSAFPEVYKAVVOMMERTPIQKWPVAVVE
YHDGSHAVIGLDRKRRESELRVIAQVYKAGSVASTSEVLYLPIKSLDPEAS
RIALLMASQEPQKMSRNFYVQGLKKKRVYLPVIGPHANIKOITIEKQAPENK
AFVLSVDELEOODLIVSYCLDAPAPPTLPVPAQVYVGGILGYSYTIQKRS
MVLVAEVLKGSOKIGADENRSKEPEEYVLOHMDQDSLHVLYQYSMTVAYEP
SEASOKDILIDRVIRIQCGRNRTNCTALRTISHSFTVSGHREBAPLVYVIGN
PASDEIKHLDIOVPLIGVGNANVOPLERTGPNPILIQDETLPRAPDLVLR
CCSGGLDIPPLSPAPDQPLVDILLDSSSPAPPEMKSFAKATSKANIGPR
LTVSVLYQVGSITTVPMVNVPEKAILLSIVDMEGEGSPICGALAGAVYLTSE
MHGAPVASKAVVILVTVSVSYVAAADARSKRVTPFEGIDRDAQVLLIAGP
AGDSNVVYLRILEDPLMVTIGNSFLKIKSGFVLCMDGNSKRPGRVMTIPDGH
VTQCPQGGILLKTHRYNODGILKPSNPSQVKKVYELDQGMKILQVYVNSIPIH
VTFDQNRKLITGSCSYLPQKRPDILVILLNAGSGAGAGQMKSTFKHSLVEL
HSDKRVTVNPIVSPVYVGGMEVNVGAILMHEVRNHLHITFTTPONNEFLQISP
KTFASKTYGJGICDENANDFMLDGTVOENTLQVETVOPQDQPLLEBQCL
VPSDSHQVJLLPLFARCHKVLAPEVYALCOQDSHQEGVAVIASYATGNTNCL
VDMRTPECASTSPSTVYNNFHHQTPHHNGVNSQHHSHQSPKSPKAVLESG
VPEACIQCIQEGDVQHOFLAWYVHOHQDITCISGRKVNLTQPPAKAPKRL
CEVALRQNDQCCPRYECYDPSYCLTPYPRCHERTLQGTITNPPCPPTGAKRK
EACKVSPSPPHPIPIRTGNTTPYFACNPNVSTVSPVSLASTATNINNTYIT
TTCLEKQVHARSTIYVQGFMEGTLVCTCTMEDAVMLRAQDSOKPCEDSCSG
FTYVJHGEQDGRCLPSACEVYITSPRGDSQSSVSGVMSAPENDCLINCEVKE
EFTVIOGHNVSPQLEEVVPSGFLSKISACDPSGCFEMALMNGVYIISQKIVM
IDVCTTGQMGVGVISGPKLEKRTKTCNCPGSGYKRENNTGCGGCGCTACTGDR
GGQIMTLKRPETLDDYDTHCKNHEGEYFMEKRYGCGPEPBHGLAEGGTAKIP
CTCTCTTCTPCTPNTITATLQYVKGSKSEVFYVITVQGRASKAMYSIDINVDQ
CSCSPRTTSPMOVALHCTGNSVYVHVLAMECKSPKRSK"
misc_feature
100..165
2389..8538
166..8533
/note="VWF pro polypeptide (AA 1-2791)"
BASE COUNT 1759 a 2467 c 2559 g 1790 t
ORIGIN

```

```

DEFINITION Sequence 3 from Patient Wt 8606096.
ACCESSION 108449
VERSION 108449.1 GI:588840
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 8588)
AUTHORS
Ginsburg D., Orkin S.H. and Kaufman, R.J.
TITLE
VON WILLEBRAND FACTOR
JOURNAL
Patent Wt 8606096 A 3 23 OCT 1986;
FEATURES
source
1..8588
/organism="unknown"
BASE COUNT 1763 a 2472 c 2559 g 1794 t
ORIGIN

```

Query Match 100.0%; Score 18; DB 6; Length 8588;
 Best Local Similarity 100.0%; Pred. No. 1,1e-02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

db 1 GTGGTCGTCGACCGCTA 18
    |||||||||||||||
db 979 GTGGTCGTCGACCGCTA 953

```

RESULT 9
 AF099154/c 8601 bp mRNA linear MAY 14 JUN 1999
 LOCUS Canis familiaris von Willebrand factor (vWF) mRNA, complete cds.
 DEFINITION AF099154
 ACCESSION AF099154
 VERSION AF099154.1 GI:4154308
 KEYWORDS
 SOURCE
 ORGANISM
 Canis familiaris.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis;
 1 (bases 1 to 8601)
 Ventura, P.J., Li, J., Yuzbasyan-Gurkan, V., Brower, G.J. and
 Schall, W.D.
 TITLE
 Complete sequence of the structural gene for canine von Willebrand
 factor and identification of a mutation causing Scottish terrier
 von Willebrand's disease
 JOURNAL
 Unpublished
 2 (bases 1 to 8601)
 Ventura, P.J., Li, J., Yuzbasyan-Gurkan, V., Brower, G.J. and
 Schall, W.D.
 TITLE
 Direct Submission
 Submitted (15-SEP-1998) Microbiology, Michigan State University,
 College of Veterinary Medicine, East Lansing, MI 48824-1314, USA
 FEATURES
 source
 1..8601
 /organism="Canis familiaris"
 /db_xref="taxon:9615"
 1..8601
 /gene="vWF"
 1..8442
 /gene="vWF"
 /note="clotting factor"
 /codon_start=1
 /product="von Willebrand factor"
 /protein_id="AA004919.1"
 /db_xref="GI:4154309"

```

Query Match 100.0%; Score 18; DB 9; Length 8575;  

Best Local Similarity 100.0%; Pred. No. 1,1e-02;  

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

db 1 GTGGTCGTCGACCGCTA 18
    |||||||||||||||
db 963 GTGGTCGTCGACCGCTA 946

```

```

RESULT 7
108449/c 108449
8588 bp INA linear PAT 02-DEC-1994
```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 3694)
AUTHORS
Montgomery, P. R., Fels, S. and Montgomery, M. W.
TITLE
Unpublished
JOURNAL
2 (bases 1 to 4994)
REFERENCE
Montgomery, P. R., Fels, S. and Montgomery, M. W.
AUTHORS
Direct Submission
TITLE
Submitted (06-AUG-1996) Blood Research Institute, The Blood Center
JOURNAL
of Southwestern Wisconsin, Box 2178, Milwaukee, WI 53201-2178, USA
FEATURES
source
1
Location/Qualifiers
1
6494

CDS

/organism="Canis familiaris"
/db_xref="taxon:9615"
/cell_type="endothelial"
105..8546
/note="vml"
/product="von Willebrand factor"
/protein_id="AAB93766.1"
/db_xref="GI:2734858"
/translation="MSPTRIKVVIIATALLPKLITKGVVPSMAKPSIFGGDFIN
TPDSMYSFAEDSYLLAGDCQEHNSVLSIGFQNGKRVSLVSLGFEDFIDLPKNTM
LQGTQSTSMFASNGSLLEAAYVYKISFAVGFVAPLTCNNEVLISSVYFNPLVY
LQGNENIFADHDFPTQGTITLSDYDFANSWALSSEORFKVPSPPSPVNSPFW
QVLMGQCQLLKASVAFARHILVDEPPVALCEPTLCTQVQMGECGGLLEAPCA
QGVVALGMDTHSYCRAPACAGMYKCYSPGRTQSLHKVLDQDQVLDQSLQHG
QLLDGHCVASAECSVAGHGYHPCATLQGTCTCRSLTMTCTNRPRTDQCLVTS
QSHKSFEDNRYFTSGICQYLLAQVQDHTFSVILEVQCADDLAVCPSTVPLPG
HNSLVKLRIGGAGVSMQGGDQIPILOGDITQHTWASVLSYGEDLDMDWGRRL
LVLTSFAVAKRTGQIATNMYNMRNDYDFPAGIALVEDEFGAMKLGACENLKH
RDPCLSLPQAFAPAFACALITSSKFPKRAVAGPQVYQNCYVDSGLDQVLS
AVANYAAGCARGVHIAIRERPCALISCPQGOVYLVQGVPCNNTCSLSYPERDCHV
CLCEGCPGLYLDERGDCVAKACPCYVGGELFQEDHIFSDHHTMCGEDGPMHCT
SGIGSLIPNLVSSPSRHSKRSLSRPMMLVLPANPARFSGLEAKTQVNTDQ
CMSTGCVSGSLCEGVNKHKCVALEKPCPHQGEJARAGETVAKIDCTQGRDKW
NCTHVCADATSAIGMAHYLTFGLKTLFEGEVNKKPKDTELVQVYQVNSNPTLRLLVGNDS
SYSPVCKKRVTLVEGDELDFEGEVNKKPKDTELVQVYQVNSNPTLRLLVGNDS
VMDRLISVTLKTYELVWGLQVNTGJUNNLTSSLSLEDEVDGNSKVMPO
CADTKRVLSSPAVGNHNNIMKQMTNSPRLTSDFGCNPLVDEPELIDGIVDT
CSGESISDGTGCDTIAVAHVCAGHKKVAMPTATGEPQNEERMLHNGVCEMPY
NSCAPACETIQGHEPLATVGRVHAKCPKGLLELLITLLEPCVYVVAAPV
RIABGKKTILNPSRHCQICHTGDNVETQACRPESLVLPTEPISSTSYEDT
PEPLHDFKSKLIDVLFVDSKSLSEDFEVLKVFVWMMHLISOKILVAVE
YHDGHAFTIKKORRPSRLRITTSQVAGSVASTSEVLTTLTLEQKIDPPAS
RIALLASQEPSRIARLVAVGLKKKVIIVYIOPHASKVILLEKAGLENK
AFVSGVDELEQRDELINYLCDLAPAPVAPVAPVAVVAGSELGVSGPKNS
MVLVVEVLGSDKIGEANFKSHREVEVIOAMDGDRIHTVLOYSYMTVEYTF
SEASKEVILQVHILVYKSNKNTNTEADYLSHSEFVSQGDREOVLYMTVEYTF
PASDEIKRMHQDIOVPIGVPHANVELEKICPMAPILIHFEMLPREADVLVQR
CTSGRQIPIVSTPTPKSPTIVVLISSSTPASYRPMKSPITKATISANIGPR
ITQVSVLYQYSITITIVMVNAVAKVHLLSVLMQVEGSPSILGALISFAVYVSE
VHGARGASAAVAVILTVDSVDSVDAEAARNSRVTPIGIDGYSVQSLSLAPG
KAGSNMRELPIELPLVATIGNSFHKLVSPVAVVEDKEDKPKVIMVTLPLCH
TVTCLDQGTLLKSHVNCDRGPPSPCNQPLPLVBRGTQVLTWTPVYVCMSSSTPIL
VTPDQGNPKLLKSGSVYLPRQNEQLLVLHAKGSKGAKCTEKETKHLGSLSEL
HSDMQVYKHYVSTPYVGGMEVNYGTMTEVRENHGLHTTPDQNEQLQULSP
RTFASKTYGLGICDENGANDFLIPLQVITVIMKALLVMTVQLKTCQVVEEPCV
VSSSHQVILSELFACHKVLAPATVYACQDPSCHKPVCAILAHALHRTKGC
VDMRANFCAMSPPSLVNHNHFGPPLCEGNTSSGQPCGCPGPNQVLEESG
VPEACQTCGSEGVHQLFETVWPHAGHQLVLTCLSGSKVNTLQDQCAVAPITGCP
CEVARLRQNAQCCPEYEVYVYVSTPVPVPEVGLQVLTLLNPGEPNNTCAKPK
DECRRESPPCPHRTYALKTKTQCCDYECACNVSVSLGLYASVAVINDQGTIT
TTCPEDVAVHRTITVYVQGFPEACVYCTCLILSVMLVLAQVSCQPEINISG
FTVYVHEGECGPRCTPSACEVYVSPVTAQSHKRVNVSIMAPSCPLINEVYKKE
EVEVQNRVSCPOLNPTVPTGQVLSKTSCEQPTCHCEQLECLNGITIGPQSLM
IDVCTIKRTVYVGVISGKLEKTKTICATVYKPKKPNQVQVQVLAITGLIP
GGQIMTLKROETIDQCDSHKCKVNEHTGATVYKPKCPPEPEHGLAEGRKIMIP
CTQCDTCEBECQDIKAKLVKAVGPKSEEDVDTYBCKKCAKAVYSIHMDVQDQ
CSCSPVITIPMVQVPRCTNOSTIYHILINAMCKSPKCKSK"

BASE COUNT
ORIGIN
1825 A 2473 C 2547 G 1849 T

Query Match 100.0%; Score 18; DB 4; Length 8694;
Best Local Similarity 100.0%; Pred. No. 1; Le-02;
Matches 18; Conserved 0; Mismatches 0; Gaps 0;
DB 968 GTGGTGCGTCCAGCGCTA 18
|||||
GTGGTGCGTCCAGCGCTA 951

RESULT 11
AR09B096/C 8802 bp DNA linear PAT 14-DEC-2001
Length
DEFINITION
Sequence 1 from patent US 6074832.
ACCESSION
AR09B096
VERSION
AR09B096.1 GI:12807353
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 8802)
AUTHORS
Verte, P. J., Brewer, C. J., Yurbasayan-Gurkan, V., Schell, M. D., and
Duffendeck, J.
TITLE
DNA encoding canine von Willebrand factor and methods of use
JOURNAL
Patent: US 6074832-A 1 13-JUN-2000;
FEATURES
source
1..8802
/organism="unknown"
BASE COUNT
1866 A 2477 C 2544 G 1845 T
ORIGIN

Query Match 100.0%; Score 18; DB 5; Length 8802;
Best Local Similarity 100.0%; Pred. No. 1; Le-02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1066 GTGGTGCGTCCAGCGCTA 18
|||||
GTGGTGCGTCCAGCGCTA 1049

RESULT 12
BD000647/C 8802 bp DNA linear PAT 31-JAN-2002
Length
DEFINITION
Polypeptide and process for producing polypeptide.
ACCESSION
BD000647
VERSION
BD000647.1 GI:18623760
KEYWORDS
JP 2000442097-A-1.
ORGANISM
Homo sapiens.
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 8805)
AUTHORS
Hans, P., Hombartz, P. C., Jovan, D. P. and Louisa, H. M. H.
TITLE
Polypeptide and process for producing polypeptide
JOURNAL
Patent: JP 2000442097 A 1 12-DEC-2000;
IMMUNO AG

COMMENT
OS Homo sapiens (human)
PN JP 2000442097-A/1
PD 12-DEC-2000
PE 18-MAY-2000 JP 2000145965
PR 01-APR-1985 NL 8500961
PI PANNAOOCK HANS, PULIVE CORNELIUS PUNNAPART, DIAHALDA PAUL, JOHAN,
PI HARTY MARGARETA HENDRIKA LOUISA
PC C12N15/009;C07K14.724;C12P21/02;C12N15/700
CC
FH KEY local:us/Qualifiers
FT source local:us/Qualifiers
FT local:us/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
1..8805
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT
1804 A 2529 C 2525 G 1847 T

1

2

517MMN: 1-5

Run on: May 27, 2003, 06:10:50 ; Search time 45.7443 Seconds

(without alignments)
11700.951 Million cell updates/sec

Title:	US-04-662-478C-23
Perfect score:	20

Sequence: 1 tctgttgggagcaatgcacaa 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DH seq length: 0
Maximum DH seq length: 2000000000
```

Post-processing: Minimum Match	08
Maximum Match	100%

Listing first 45 summaries

Datavase :

1:	qb_bar *
2:	qb_hfq *
3:	qb_in *
4:	qb_ov *
5:	qb_pat *
6:	qb_pod *
7:	qb_pl *
8:	qb_pl *
9:	qb_pro *
10:	qb_rts *
11:	qb_rts *
12:	qb_sy *
13:	qb_un *
14:	qb_vt *
15:	em_bar *
16:	em_hum *
17:	em_hum *
18:	em_in *
19:	em_mu *
20:	em_com *
21:	em_or *
22:	em_ov *
23:	em_pat *
24:	em_pl *
25:	em_pl *
26:	em_rts *
27:	em_rts *
28:	em_un *
29:	em_vt *
30:	em_hfq_hum *
31:	em_hfq_hum *
32:	em_hfq_or_bar *
33:	em_hfq_rts *
34:	em_hfq_pln *
35:	em_hfq_pod *
36:	em_hfq_rts *
37:	em_hfq_vt *
38:	em_sy *
39:	em_hfq_hum *
40:	em_hfq_rts *
41:	em_hfq_or_bar *

Prod. No. is the number of results predicted by chance to have a

score greater than 1 point for the score and is derived by analysis of the data.

517MMN: 1-5

Result	No.	Score	Category	Mat. No.	Ref.	Ill.
	1	20	1	1	1	AR-96-23
	2	20	1	1	1	AR-98-23
	3	20	1	1	1	AR-99-23
	4	20	1	1	1	AR-99-23
	5	20	1	1	1	AR-99-23
	6	20	1	1	1	AR-99-23
	7	19	1	1	1	AR-99-23
	8	19	1	1	1	AR-99-23
	9	19	1	1	1	AR-99-23
	10	19	1	1	1	AR-99-23
	11	19	1	1	1	AR-99-23
	12	19	1	1	1	AR-99-23
	13	19	1	1	1	AR-99-23
	14	19	1	1	1	AR-99-23
	15	19	1	1	1	AR-99-23
	16	19	1	1	1	AR-99-23
	17	19	1	1	1	AR-99-23
	18	18.4	1	1	1	AR-99-23
	19	18.4	1	1	1	AR-99-23
	20	18.4	1	1	1	AR-99-23
	21	18.4	1	1	1	AR-99-23
	22	18.4	1	1	1	AR-99-23
	23	18.4	1	1	1	AR-99-23
	24	18.4	1	1	1	AR-99-23
	25	18.4	1	1	1	AR-99-23
	26	18.4	1	1	1	AR-99-23
	27	18.4	1	1	1	AR-99-23
	28	18.4	1	1	1	AR-99-23
	29	18.4	1	1	1	AR-99-23
	30	18.4	1	1	1	AR-99-23
	31	18.4	1	1	1	AR-99-23
	32	18.4	1	1	1	AR-99-23
	33	18.4	1	1	1	AR-99-23
	34	18.4	1	1	1	AR-99-23
	35	18.4	1	1	1	AR-99-23
	36	18.4	1	1	1	AR-99-23
	37	18.4	1	1	1	AR-99-23
	38	18.4	1	1	1	AR-99-23
	39	18.4	1	1	1	AR-99-23
	40	18.4	1	1	1	AR-99-23
	41	18.4	1	1	1	AR-99-23
	42	18.4	1	1	1	AR-99-23
	43	18.4	1	1	1	AR-99-23
	44	18.4	1	1	1	AR-99-23
	45	18.4	1	1	1	AR-99-23

AI.1.1 NM+NI

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
ARO98114	ARO98114	Sequence of the patent US 6,104,673	ARO98114	ARO98114.1	11.1.800471	Unknt. acc.	Unknt. acc.	1 (bases) of 2	Ventri, E. and Brown, G. A. Yuzbasova	DNA encoding the gene for the wild type of the	Patent US 6,104,673

KAGSNMVLRLIEDLPIVATLIGNSFEHLKSGEDRCVDEDEGMEKRRGVDVWLLPDDCH
 VTGCLPFGGILLKSHRYNCDRGPPSPSPNUPPLKAVETETNHWIRPVVMSNSTHPI
 VTFDQUNKJTGSCSYLFPKEDDLEVLHNGSGSGAETOMKSTIEVHGLSEL
 HSDMONTNRLVSI PVYGDMEVNVGTIMYEFREHILHETFPDNNELQSLSP
 RFPASKTVLGGITDENANFII RPSIVITDKALLIEMVITLQKTCVYEDUC
 VSSSHQCVJISELFAECHKAVATPYAMCOPDSCHKKVCALAIYALACTKGC
 VDMRANCGMSPSPSLVYNGEHCGLPCLDGNISGHPSPSGDCHPNUVMELESG
 VPEFACGQISDEGVHRQPLFEVWFAHQDQVITICISPKVNCILQCPPLAVALTECP
 CEVALRPNANQCCPPEYCVYDIYSCILPEYVEDELOMTLLNPGVYRHNPTAKRK
 DECRHSPSPGPHRTPALRKTQCCDDEYACNCSVSTVSLGLYLASAVTNGCCTT
 TTCPEDKVCVHRGIIYVQGFWEFACVCTCIULEDSVMGLKVAJCSUKPCEDGCTT
 FTYVILHEDGDCGCLPSACEVYIOSPGRDASQHWKNGVSHASPDNPCLINCEVKE
 EYVQORVASCPOLVNPTCPGHPOLSKTSECCCHCEPLALNLNGHIIIGKSLM
 IDVCTTCCTVQGVISAPKIFPKITGCAPIGLQYKPKKMGEGCGHOLPIACTGLP
 GGOITLKRDETTIGDCTSHCKYKNGHGEVIMKRVYRPPPHFKCTAPGSKIMIP
 CTCDCICEPECKDITAKLQRYKVGDKSEEDVIRCEGCKASKAVYSTIHMEDVDD
 CSCSPDTOTPMOYPLKCTNSLTYHEILNMACRCSPRKSK*

BASE COUNT 1825 a 2473 c 2547 g 1849 t

ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 8694;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCTGAGAGTGGCCAG 20
 ||||||||||||||||
 DB 763 TCTCTGAGAGTGGCCAG 782

RESULT 5
 LOCUS AR098096 8802 bp DNA linear PAT 14-FEB-2001
 DEFINITION Sequence 1 from patent US 6074832.
 ACCESSION AR098096
 VERSION AR098096.1 GI:12807353
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 1 (bases 1 to 8802)
 AUTHORS Venla P T., Bremer, O. J., Yuzhaslyan-Gurkan, V., Schall, W.D. and
 Buftdeck, J.
 TITLE DNA encoding canine von Willebrand factor and methods of use
 JOURNAL Patent: US 6074832-A 1-13-JUN-2000;
 FEATURES
 source location/qualifiers
 1..8802
 /organism="unknown"

BASE COUNT 1886 a 2477 c 2544 g 1895 t

ORIGIN

Query Match 100.0%; Score 20; PP 6; Length 8802;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCTGAGAGTGGCCAG 20
 ||||||||||||||||
 DB 861 TCTCTGAGAGTGGCCAG 980

RESULT 7
 LOCUS HUMWVFA04 734 bp DNA linear PRI 14-JAN-1995
 DEFINITION Human von Willebrand factor gene, exon 7.
 ACCESSION M25831 M25716
 VERSION M25831.1 GI:34032C
 KEYWORDS
 SPMENT
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

4 of 38
 Human placenta, leukocyte, fetal liver and cell line PB19 DNA
 clones lambda- [1,2,7,9,11], h[6,7,18], c[11,2,3,4,5,9,14,18] and
 pME[5,8,123].

REFERENCE 1 (bases 1 to 734)
 AUTHORS Mannes, D. J., Tukey, R. A., Westfield, L. A., Worrall, N. K.,
 Shelton-Idol, B. B., Sorace, J. M., Avery, Y. G. and Sadler, J. E.
 TITLE Structure of the gene for human von Willebrand factor
 JOURNAL J. Biol. Chem. 264 (35), 19514-19527 (1989)
 MEDLINE 90062044
 PUBMED 2564182
 COMMENT Draft entry and computer-readable sequence for [1] kindly provided
 by J.E. Sadler, 30-JUN-1989.
 location/qualifiers
 1..734
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="12pctcr.p12"
 <1..108
 /gene="VWF"
 /number=6
 109..325
 /gene="VWF"
 /number=7
 326..734
 /gene="VWF"

BASE COUNT 106 a 247 c 216 g 165 t
 ORIGIN About 13.7 kb after segment 3, chromosome 12pter.p12.

Query Match 95.0%; Score 19; DB 9; Length 734;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCTGTGGAGCAGTGCAG 20
 ||||||||||||||||
 DB 111 CCTGTGGAGCAGTGCAG 129

RESULT 8
 LOCUS AY004876 1110 bp mRNA linear MAM 14-JUL-2001
 DEFINITION Sus scrofa von Willebrand Factor precursor (VWF) mRNA, partial cds.
 ACCESSION AY004876
 VERSION AY004876.1 GI:14718399
 KEYWORDS
 SOURCE Sus scrofa.
 ORGANISM Sus scrofa.
 1 (bases 1 to 1110)
 AUTHORS Fahs, S.A., Nichols, T.C., Seaman, W.T. and Montgomery, R.K.
 TITLE Porcine VWF cDNA - 5' UT and 1st 1032 bp of propeptide
 JOURNAL Unpublished
 2 (bases 1 to 1110)
 REFERENCE Fahs, S.A., Nichols, T.C., Seaman, W.T. and Montgomery, R.K.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2000) VWF Biology, Blood Research Institute, P.O.
 Box 2178, Milwaukee, WI 53201 2178, USA
 location/qualifiers
 1..1110
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /chromosome="5"
 /map="5q21"
 /cell_type="platelet"
 1..>1110
 /gene="VWF"
 1..78
 /gene="VWF"
 79..>1110
 /gene="VWF"
 /note="regulation protein; preproprotein"
 /codon_start=1
 /product="von Willebrand Factor precursor"
 /protein_id="AA09092.1"
 /db_xref="GI:14718400"

```

      sig_peptide      79..144
      misc_feature      /gene="WVF"
                        /note="Peptides von Willebrand factor preproprotein; vW Ag1"
BASE COUNT      212 a      329 c      337 g      232 t
ORIGIN
Query Match      95.0%; Score 19; DB 4; length 1110;
Best local similarity 100.0%; Pred No. 62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY      2 CCTGTGAAATATATTTAAATTTT
      745 CCTGTGAAATATATTTAAATTTT

RESULT 9
LOCUS      HSWF8
DEFINITION      Human mRNA: treatment (5' terminus) for von Willebrand factor (WVF)
VERSION      X04146.1 GI:37941
KEYWORDS      cDNA; protein; plasma; protein; signal peptide; von Willebrand factor.
SUBJECT      Homo sapiens.
ORGANISM      Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE      1 (bases 1 to 4429)
AUTHORS      Verweij, C.L., Diegdaarde, P.J., Hart, M. and Pannekoek, H.
TITLE      Full-length von Willebrand factor (WVF) cDNA encodes a highly repetitive protein considerably larger than the mature WVF subunit.
JOURNAL      EMBO J. 5 (4) 1849-1847 (1986)
MEDLINE      87004550
COMMENT      see also HSWF2
FEATURES
     source          1..4429
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /protein_id="CAA27765.1"
                        /translation="MIPARFVLAIAIIFPTICARTGPRSTAGSIFGSDPVN
                        TFGDSHTFAGCGSTLAAGCQKSHSTIGDQNKRVLSVYLGEFFDILFVNGIV
                        TQTPGVSPVASKGLYLETAGVYKLSGAVGPAKIDGSCNFQVLLSDYFNKTC
                        LGCNENIFAADIPMTQSGITSDPYDANSMALSSGQWCRASPSSCNLSGDEM
                        KGLMEQCQLKSTVFARCIPIVDEPVALCEKILCEAGGLECACALLEVARTCA
                        QRMVLYGWTHTSASPVTPAEMPEKGVSPAPFTQSLHINEMQDEPGVSCPTG
                        QLLDEGLVESTFGCVSHSKRPPTSTSPQNTGCPNSQWQNSNFGPGGPTLQ
                        QSHKSPRRVFSTGQYLLAKTQGHSPSTVETVGTADRRPAVTSVATPIHIS
                        LNSLKLKTRACVAMQGVDPPIIKKPIQPIVASVPLSTGEDLQMLQDQCPIL
                        LVKISVYAGKTQITQNYNNGNQHETPSLAEVPELQNMKLLHGDQDLQKH
                        SDPCALNPMTPFSEBAVATSPFFAPRAVSPPIYLINCPYDVGSGDGECLTG
                        ALASYAAACAGVAVAMPPEPPEPIINPKGAVLQCTPCLTQPSLSTYDEGCNA
                        CLEGGCPGLTMDERQVKAQCPYDGLFQPEDIFSDHDMCGCEGEMHIC
                        SGVPSLLDPAVLSPISHRSKRSLSGRPWKLVCPAPNLRACITKCTQCNVLE
                        CSMGCVSCGLCPGVMHENVALFPGCPHQKTPAPSETYKIGTQVCGRRK
                        NCLDHVCAVTSITGMARLTDLKLTLPQGVVLVDQDQSGSPGFRLLVGRKQ
                        SHPSVCKRPVITLVEGELFELQGVAVPKMKDHFEEVSESRVITLLGLVALSV
                        WMQRLISVVKIQTYSKRVKQDINPQINNDLISNLVEEDPVDQFGSMKVVSSQ

```

```

      sig_peptide      213..4429
      misc_feature      /gene="WVF"
                        /note="Peptides von Willebrand factor preproprotein; vW Ag1"
BASE COUNT      212 a      329 c      337 g      232 t
ORIGIN
Query Match      95.0%; Score 19; DB 4; length 1110;
Best local similarity 100.0%; Pred No. 62;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY      2 CCTGTGAAATATATTTAAATTTT
      745 CCTGTGAAATATATTTAAATTTT

RESULT 10
LOCUS      A09245
DEFINITION      Partial synthetic human von Willebrand factor
VERSION      A09245
KEYWORDS      A09245
SUBJECT      Homo sapiens.
ORGANISM      Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE      1 (bases 1 to 720)
AUTHORS      Funderud, H., Verweij, C.L., Diegdaarde, P.J., Hart, M. and Pannekoek, H.
TITLE      Preparation of the human von Willebrand factor (vWF) cDNA.
JOURNAL      STICHING VAN DE VWA 14 511-511
MEDLINE      87004550
COMMENT      see also HSWF2
FEATURES
     source          1..720
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /protein_id="A09245.1"
                        /translation="MIPARFVLAIAIIFPTICARTGPRSTAGSIFGSDPVN
                        TFGDSHTFAGCGSTLAAGCQKSHSTIGDQNKRVLSVYLGEFFDILFVNGIV
                        TQTPGVSPVASKGLYLETAGVYKLSGAVGPAKIDGSCNFQVLLSDYFNKTC
                        LGCNENIFAADIPMTQSGITSDPYDANSMALSSGQWCRASPSSCNLSGDEM
                        KGLMEQCQLKSTVFARCIPIVDEPVALCEKILCEAGGLECACALLEVARTCA
                        QRMVLYGWTHTSASPVTPAEMPEKGVSPAPFTQSLHINEMQDEPGVSCPTG
                        QLLDEGLVESTFGCVSHSKRPPTSTSPQNTGCPNSQWQNSNFGPGGPTLQ
                        QSHKSPRRVFSTGQYLLAKTQGHSPSTVETVGTADRRPAVTSVATPIHIS
                        LNSLKLKTRACVAMQGVDPPIIKKPIQPIVASVPLSTGEDLQMLQDQCPIL
                        LVKISVYAGKTQITQNYNNGNQHETPSLAEVPELQNMKLLHGDQDLQKH
                        SDPCALNPMTPFSEBAVATSPFFAPRAVSPPIYLINCPYDVGSGDGECLTG
                        ALASYAAACAGVAVAMPPEPPEPIINPKGAVLQCTPCLTQPSLSTYDEGCNA
                        CLEGGCPGLTMDERQVKAQCPYDGLFQPEDIFSDHDMCGCEGEMHIC
                        SGVPSLLDPAVLSPISHRSKRSLSGRPWKLVCPAPNLRACITKCTQCNVLE
                        CSMGCVSCGLCPGVMHENVALFPGCPHQKTPAPSETYKIGTQVCGRRK
                        NCLDHVCAVTSITGMARLTDLKLTLPQGVVLVDQDQSGSPGFRLLVGRKQ
                        SHPSVCKRPVITLVEGELFELQGVAVPKMKDHFEEVSESRVITLLGLVALSV
                        WMQRLISVVKIQTYSKRVKQDINPQINNDLISNLVEEDPVDQFGSMKVVSSQ

```


RESULT 12

LOCUS

108449

Sequence 3 from Patent WO 8606096

Accession 108449

Version 108449.1

Keywords G1:588840

Source

Organism

Unidentified

Reference

1 (bases 1 to 8588)

Authors Ginsburg, D., Orkin, S. H. and Kaufman, R. J.

Title WGN WILDBRAND PATENT

Journal Patent: WO 8606096-A 3 23-OCT-1986

Features Location/Qualifiers

Source

1..8588

/organism="Homo sapiens"

BASE COUNT

1763 A

2472 C

2559 G

1794 T

Query Match

Best Local Similarity

Matches

1% Conservative

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8588

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

RESULT 14

LOCUS

E01056

CDNA sequence from library of human

Accession E01056

Version E01056.1

Keywords JP 1997-000000 A/1

Source

Organism

Homo sapiens

Reference

1 (bases 1 to 8805)

Authors Ginsburg, D., Orkin, S. H. and Kaufman, R. J.

Title WGN WILDBRAND PATENT

Journal Patent: WO 8606096-A 3 23-OCT-1986

Features Location/Qualifiers

Source

1..8805

/organism="Homo sapiens"

BASE COUNT

1763 A

2472 C

2559 G

1794 T

Query Match

Best Local Similarity

Matches

1% Conservative

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

Score 192

DB 6: Length 8805

Pred. No. 47

0% Mismatches

0% Indels

0% Gaps

0%

REFERENCE
AUTHORS

Mammalia: Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 111904)

Murphy, A. R., Ardura, A. D., Ruck, J., Buzza, C., Chen, Z., Ding, Y., Dugan, S., Durbin, J., Forcum, J., Garcia, C., Gorrell, J. H., Gorrell, L. L., Hernandez, J., Issari, A., Jackson, L., Kiehl, S., Kondejowski, N., Lai, S., Leal, B., Lee, E., Lichare, O., Liu, W., Logan, O., Lu, J., Marandel, L., Martinez, C., Merscher, S., Miller, A., Montgomery, K., Oswal, G., Pampell, L. K., Parish, B. J., Petz, J., Rashid, N. D., Rives, C., Scherer, S. E., Shen, H., Shim, C., Simon, M., Vo, Q., Williamson, A., Worley, K. C., Xiang, A. M., Yang, R., Yu, W., Zhou, X., Kuchelapali, R., Nelson, D., and Gibbs, R. A.

REFERENCE
AUTHORS

2 (bases 1 to 111904)

REFERENCE
AUTHORS

3 (bases 1 to 111904)

REFERENCE
AUTHORS

4 (bases 1 to 111904)

REFERENCE
AUTHORS

5 (bases 1 to 111904)

REFERENCE
AUTHORS

6 (bases 1 to 111904)

REFERENCE
AUTHORS

7 (bases 1 to 111904)

REFERENCE
AUTHORS

8 (bases 1 to 111904)

REFERENCE
AUTHORS

9 (bases 1 to 111904)

REFERENCE
AUTHORS

10 (bases 1 to 111904)

REFERENCE
AUTHORS

11 (bases 1 to 111904)

REFERENCE
AUTHORS

12 (bases 1 to 111904)

REFERENCE
AUTHORS

13 (bases 1 to 111904)

REFERENCE
AUTHORS

14 (bases 1 to 111904)

REFERENCE
AUTHORS

15 (bases 1 to 111904)

REFERENCE
AUTHORS

16 (bases 1 to 111904)

REFERENCE
AUTHORS

17 (bases 1 to 111904)

REFERENCE
AUTHORS

18 (bases 1 to 111904)

REFERENCE
AUTHORS

19 (bases 1 to 111904)

REFERENCE
AUTHORS

20 (bases 1 to 111904)

REFERENCE
AUTHORS

21 (bases 1 to 111904)

REFERENCE
AUTHORS

22 (bases 1 to 111904)

REFERENCE
AUTHORS

23 (bases 1 to 111904)

REFERENCE
AUTHORS

24 (bases 1 to 111904)

REFERENCE
AUTHORS

25 (bases 1 to 111904)

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards. Estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.ncbi.nlm.nih.gov/qual/qualinfo/qualbank.annotation.html>.

QUALSTAT-REPORT

Summary Statistics

Contig length:	94462
Phrap values in estimate:	93326
Average error rate (BCM-Phrap estimate):	0.00022054
Fraction of Phrap values less than 40:	0.045195
Number of consensus changing edits:	3
Number of N's in consensus:	0

Consensus changing edits

Position	Original/Context	Edited/Context
19379	agagcttaaa(g)acacggttgc	agagcttaaa(g)acacggttgc
29991	tttaagtaga(a)atagagtttc	tttaagtaga(a)atagagtttc
89256	gcattgggtct(t)cttggccca	gcattgggtct(t)cttggccca

Distribution of Quality < 40 Bases

# bases	5	10	15	20	25	30	35	40
10001								
9001								
8001								
7001								
6001								
5001								
4001								
3001								
2001								
1001								
01								

FEATURES

source

Version: 1.01 qxto.

location/qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="12"

/clone="RPC13-467F14"

43.319

/rpl_family="A1uSx"

complement(335..441)

/rpl_family="MER1A"

complement(783..1081)

/rpl_family="A1uSx"

complement(1109..1249)

/rpl_family="12"

complement(1250..1412)

/rpl_family="MER5A"

1962..2258

/rpl_family="A1uSx"

complement(2816..3008)

/rpl_family="A1uSp"

4761..4911

/rpl_family="MER20"

5307..5504

/rpl_family="A1uSx"

5990..6043

/rpl_family="MER"

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect = 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

```

repeat_region      complement(5942, .7020)
                    /rpl_family="LIM4"
repeat_region      7053, .7347
                    /rpl_family="AluY"
repeat_region      complement(7736, .7841)
                    /rpl_family="(A)n"
repeat_region      complement(8668, .8976)
                    /rpl_family="AluY"
repeat_region      9030, .9259
                    /rpl_family="AluB"
repeat_region      complement(9835, .9910)
                    /rpl_family="MIR"
repeat_region      10696, .10991
                    /rpl_family="AluY"
repeat_region      complement(11228, .11294)
                    /rpl_family="LIM4"
repeat_region      complement(11428, .11542)
                    /rpl_family="SLAMF"
repeat_region      complement(11646, .11945)
                    /rpl_family="AluY"
repeat_region      complement(12278, .12316)
                    /rpl_family="(CA)n"
repeat_region      complement(12312, .12633)
                    /rpl_family="AluX"
misc_feature        12677, .12817
                    /function="low coverage"
repeat_region      12688, .12674
                    /rpl_family="AluY"
repeat_region      complement(13020, .13203)
                    /rpl_family="LIM4"
repeat_region      complement(13222, .13689)
                    /rpl_family="LIME1"
repeat_region      complement(13750, .14053)
                    /rpl_family="AluS4"
repeat_region      complement(14054, .14412)
                    /rpl_family="LIM3"
repeat_region      complement(14440, .14585)
                    /rpl_family="LIM3"
repeat_region      complement(14586, .14889)
                    /rpl_family="AluS6"
misc_feature        14757, .14787
                    /function="low coverage"
repeat_region      15110, .15408
                    /rpl_family="AluX"

```

```

Query Match      95.0% Score 19; DB 9; length 11904;
Best local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
UY      2 CTTGTGGAGCAATGCCAG 20
        |||
Db 45096 CTTGTGGAGCAATGCCAG 45114

```

Search completed: May 27, 2003, 07:17:17
 Job time : 75.0777 secs

